



Universiteit
Leiden
The Netherlands

Image analysis for gene expression based phenotype characterization in yeast cells

Tleis, M.

Citation

Tleis, M. (2016, July 6). *Image analysis for gene expression based phenotype characterization in yeast cells*. Retrieved from <https://hdl.handle.net/1887/41480>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/41480>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/41480> holds various files of this Leiden University dissertation

Author: Tleis, Mohamed

Title: Image analysis for gene expression based phenotype characterization in yeast cells

Issue Date: 2016-07-06

Image Analysis for Gene Expression based
Phenotype Characterization in Yeast Cells

Mohamed Tleis

July 2016

ISBN: 978-94-028-0236-8

Copyright © by Mohamed S. Tleis. All rights reserved. No part of this thesis may be reproduced or transmitted in any form, by any means, electronic or mechanical, without prior written permission from the author.

Image on the back cover is from a GFP yeast cell-line expressing Bmh1 made with Zeiss Confocal Microscope.

This research was partly supported by:

- *Erasmus Mundus JoSyLeEN program*
- *Raymond-Sackler organisation.*
- *Landelijke Stichting voor Blinden en Slechthzienden (LSBS) organisation.*



Image Analysis for Gene Expression based Phenotype Characterization in Yeast Cells

Proefschrift

te verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op graag van Rector Magnificus Prof. Mr. C.J.J.M Stolker

volgens besluit van het College voor Promoties

te verdedigen op woensdag 6 juli 2016

klokke 10.00 uur

door

Mohamed Tleis

geboren te Brital, Libanon, in 1982

Promotiecommissie

Promotoren: Prof. dr. T.H.W. Bäck
Prof. dr. A. Plaat

Co-Promotor: Dr. ir. F.J. Verbeek

Overige leden: Prof. dr. H. Blockeel Katholieke Universiteit Leuven, Belgium
Prof. dr. T. Gevers University of Amsterdam
Dr. G.P.H van Heusden

Dedicated to humanity ...

Contents

List of Figures	v
List of Tables	vii
1 Introduction	1
1.1 Research Objective	2
1.2 Cytomics and <i>Saccharomyces cerevisiae</i>	2
1.2.1 Gene Expression and Measurement	2
1.2.2 Fluorescent Microscopy and Flow Cytometry	4
1.2.3 Cytomics and Fluorescent Proteins	5
1.2.4 <i>Saccharomyces cerevisiae</i> in cytomics	7
1.3 Pattern Recognition	8
1.3.1 Image Acquisition	8
1.3.2 Image Analysis	10
1.3.3 Data Analysis	13
1.4 Thesis Structure	13
2 Yeast Analysis Platform	17
2.1 Background	18
2.2 YeastAnalysis Workflow	19
2.3 Segmentation Module	19
2.3.1 Filter-Based Methods	19
2.3.2 Hough Transform and Minimal Path	23
2.4 Measurement Module	24
2.4.1 First order statistics based features	25
2.4.2 Texture Measurement	25
2.5 Data Analysis and Visualization Module	28
2.6 System GUI	29
2.7 Workflow validation	31
2.8 Development of <i>YeastAnalysis</i>	32
2.9 Conclusion	32

3	Hough-based Contour Extraction and Optimization	33
3.1	Background	34
3.2	Hough Transform	34
3.2.1	Filling the Accumulator	35
3.2.2	Accumulator Threshold	35
3.3	Polar Transformation	37
3.3.1	Polar Coordinate system	37
3.3.2	Cartesian to Polar System	37
3.3.3	Polar Image of Yeast Cells	39
3.4	Minimal Path Algorithms	39
3.4.1	Grey-weighted Distance Transform	40
3.4.2	Circular Shortest Path	40
3.5	Contour Optimization	44
3.5.1	Background	45
3.5.2	Control parameters	46
3.5.3	The Expansion Algorithm	47
3.6	Validation	50
3.6.1	Pratt Score	52
3.6.2	F1-Measure	53
3.6.3	Results	53
3.6.4	Robustness under Noise	56
3.7	Conclusion	58
4	Machine Learning to Identify Subtle Patterns and Improve Object Recognition	59
4.1	Introduction	60
4.2	Sophisticated Features	61
4.2.1	Feature Extraction Techniques	62
4.3	Constructing the Classification Model	66
4.3.1	Imbalanced Dataset and Sampling Techniques	66
4.3.2	Feature Scaling or Normalization	68
4.3.3	Feature Selection	69
4.3.4	Building a Classifier	70
4.3.5	Evaluation metrics, <i>ROC</i> and <i>AUC</i>	70
4.3.6	Classifiers Evaluated	71
4.4	Results	71
4.4.1	Power of Sampling	73
4.4.2	The effect of Normalization and Feature Selection	74
4.4.3	The Powerful discriminators	75
4.4.4	Feature sets performance	79
4.5	Discussion	81

5	Role of 14-3-3 proteins and Nha1 antiporter in the response of <i>S.cerevisiae</i> to salt stress	83
5.1	Introduction	84
5.2	Materials and Methods	84
5.2.1	Yeast strains and cultivation	85
5.2.2	Confocal microscopy and flow cytometry	86
5.3	Experiment Analysis	86
5.3.1	Segmentation	86
5.3.2	Measurement	87
5.3.3	Data Analysis	87
5.4	Results	90
5.4.1	Yeast Vacuoles	96
5.5	Conclusion	102
6	Discussion	103
6.1	Research Problem and Questions	104
6.2	Image Analysis Pipeline	104
6.3	Ovoid Objects Segmentation	106
6.4	Machine Learning and Feature Sets	108
6.5	Image Based Proteomics	110
	Bibliography	111
	Summary	125
	Samenvatting	127
	List of Publications	129
	About the Author	131

List of Figures

2.1	Workflow of the <i>YeastAnalysis</i> Platform.	20
2.2	Applying Sigma Filter on a Fluorescence Image.	21
2.3	Median Filter.	22
2.4	Triangle auto threshold.	22
2.5	Threshold Image and Separating Cells.	24
2.6	Graph Charts to visualize measurement.	29
2.7	<i>GUI</i> of the Yeast Analysis platform.	30
2.8	Sample report from the experiment analysis.	31
2.9	Results of the Flow Cytometry test.	31
3.1	The Hough Accumulator.	35
3.2	Filling the Hough Accumulator.	36
3.3	Extracting the Circles from Image.	37
3.4	Geometrical interpretation of relationship between polar and cartesian coordinates.	38
3.5	Polar transform of a yeast cell image.	39
3.6	Hough transform and minimal path algorithm	41
3.7	Hough transform and grey-weighted distance transform	43
3.8	Finding Circular Shortest Path	44
3.9	Sample application of circular shortest path on <i>S. cerevisiae</i> cell.	45
3.10	Flow chart explaining the expansion algorithm.	48
3.11	Expanding the Initial estimate of the contour.	51
3.12	Initial vs. expanded contour	52
3.13	Number of cells with higher Pratt score (wins).	55
3.14	Optimal Segmentation Parameters	57
3.15	F1-score at various noise levels.	58
4.1	Workflow for building the classification models.	61
4.2	<i>S. cerevisiae</i> Yeast Cell Images.	62
4.3	<i>AUC</i> and $A_{\min}$ of classifiers for raw dataset S_1	73
4.4	<i>AUC</i> and $A_{\min}$ of classifiers for raw dataset S_2	74
4.5	$A_{\min}$ of classifiers for raw and sampled dataset S_1	77
4.6	$A_{\min}$ of classifiers for raw and sampled dataset S_2	77

4.7	$A_{\min}$ of classifiers for raw, sampled and scaled dataset S_1	78
4.8	$A_{\min}$ of classifiers for raw, sampled and scaled dataset S_2	78
4.9	$A_{\min}$ of <i>SMO</i> classifier for raw dataset S_1	79
4.10	<i>ROC</i> analysis and <i>AUC</i> value using various feature sets	80
4.11	Performance of C4.5 and Logistic classifiers using moment invariant features.	80
5.1	Confocal laser scanning microscopy images	85
5.2	Detecting contours by Hough Transform and Minimal Path algorithm.	88
5.3	Overlay on a bright-field image.	89
5.4	Automatically generated overlay of estimated cell membrane locations.	91
5.5	Visualization of data generated by <i>YeastAnalysis</i>	93
5.6	Membrane Intensity of $\Delta bmh1 Nha1$ -GFP cells.	94
5.7	Flow cytometric analysis of the effect of 0.5 M NaCl.	95
5.8	Central vacuole size in <i>BMH1</i> -GFP cell strain.	96
5.9	Central Vacuole size in <i>BMH1</i> -GFP after outliers removal.	98
5.10	Vacuoles as dark spots in the yeast cell.	98
5.11	Number of Vacuoles under salt stress.	99
5.12	Total vacuole sizes in <i>BMH1</i> -GFP	99
5.13	Total vacuole sizes in <i>BMH1</i> -GFP after outliers removal.	100
6.1	Model for further platform development	107

List of Tables

2.1	Features based on first order statistics.	26
2.2	Basic texture measurement.	27
3.1	Detection Rates and F1-measure on artificial dataset.	54
3.2	Detection Rates and F1-measure on yeast images.	54
3.3	Comparing Detected Contours of yeast images.	54
3.4	F1-measure and average Pratt Score of different segmentation methods	55
3.5	Accuracy of detected contours.	55
4.1	Co-occurrence-matrix based features	65
4.2	Classification Algorithms evaluated in this study	72
4.3	AUC , A_{min} and ACC of classification algorithms using raw dataset S_1	75
4.4	AUC and Accuracy of the top 10 classification algorithms using raw dataset S_2	76
5.1	Yeast strains used in this study.	86
5.2	Interpretation of texture measurements in yeast.	90
5.3	Cell size and GFP fluorescence using <i>YeastAnalysis</i>	92
5.4	GFP fluorescence using flow cytometry	92
5.5	Analysis of vacuole size in Bmh1 14-3-3 protein	97
5.6	Vacuoles Size and number of vacuoles experiments repeated three additional times	101
5.7	Vacuoles Size and number of vacuoles experiments repeated three additional times after outliers removal	101

